Journal of Memory and Language **40,** 577–592 (1999)
Article ID jmla.1999.2634, available online at http://www.idealibrary.com on IDEAL®

# Offering a Hand to Pragmatic Understanding: The Role of Speech and Gesture in Comprehension and Memory

Spencer D. Kelly and Dale J. Barr

*The University of Chicago*

and

R. Breckinridge Church and Katheryn Lynch

*Northeastern Illinois University*

Most theories of pragmatics take as the basic unit of communication the verbal content of spoken or written utterances. However, many of these theories have overlooked the fact that important information about an utterance's meaning can be conveyed nonverbally. In the present study, we investigate the pragmatic role that hand gestures play in language comprehension and memory. In Experiments 1 and 2, we found that people were more likely to interpret an utterance as an indirect request when speech was accompanied by a relevant pointing gesture than when speech or gesture was presented alone. Following up on this, Experiment 3 supported the idea that speech and gesture mutually disambiguate the meanings of one another. Finally, Experiment 4 generalized the findings to different types of speech acts (recollection of events) with a different type of gesture (iconic gestures). The results from these experiments suggest that broader units of analysis beyond the verbal message may be needed in studying pragmatic understanding. 

It is a common observation that what people say is often quite different from what they *mean.* For instance, when someone says, "It's getting hot in here," the speaker is usually doing more than just merely commenting on the temperature. The speaker might be, for example, requesting that someone open a window, reproaching a roommate who forgot to turn down the thermostat, or proposing to change the topic of conversation. The fact that the same utterance can do many different things poses an important question for language comprehension: What information can we use to know how an utterance is intended?

A long tradition of research in pragmatics has sought to determine what kinds of contextual information can disambiguate pragmatic meaning. This tradition is characterized by the assumption that we can infer what people mean based on knowledge concerning kinds of speech acts (Austin, 1962; Searle, 1975), conventions governing the flow of discourse (Grice, 1975), and common ground or "mutual knowledge" (Clark & Marshall, 1981). This set of assumptions has been extremely influential on research concerning language comprehension. Note, however, that it presupposes that information about a speaker's intention lies somewhere outside of what is communicated—which, in turn, is traditionally presumed to be equivalent to the *spoken* portion of the message. In fact, the very labels that are almost universally attached to

This study reflects an equal contribution from the first two authors—the order in which their two names appear was determined randomly. We thank the numerous students and faculty from the Psychology Department at the University of Chicago who provided helpful ideas and criticism, especially Susan Goldin-Meadow, Sid Horton, Boaz Keysar, Sotaro Kita, Howard Nusbaum, and Tom Trabasso.

Address correspondence and reprint requests to Spencer D. Kelly, Department of Psychology, The University of Chicago, 5848 South University Avenue, Chicago, IL 60637. E-mail: kelly@ccp.uchicago.edu.



0749-596X/99 $30.00

participants in the communication process attest to this point: The terms "speaker" and "listener" suggest that the fundamental piece of information in communication is *verbal.*

However, in face-to-face interaction we do more than just speak: We glance. We point. We smile. The voice is but one component of a larger system of bodily expression. Through our facial expressions we can express pleasure or distaste; by pointing or looking at objects in the environment, we can direct another's attention to them. All of these nonverbal (gestural) behaviors add important information to a communicator's speech.[1] Information about a communicator's intention may be conveyed by these behaviors and this, in turn, could make the meaning of the utterance more clear (Argyle, 1973; Baldwin, 1993; Bruner, 1984; McNeill, 1992; Tomasello, 1992). Because these nonverbal behaviors are good visual cues to a communicator's intention, it seems likely that they would be useful to addressees when they interpret pragmatically ambiguous utterances.

One type of nonverbal behavior that has received much attention recently is hand gesture. Focusing primarily on language production, David McNeill and colleagues (1985, 1987, 1992) note that many different types of hand gestures are ubiquitous in face-to-face communication and occur simultaneously with speech. Communicators produce gestural movements of their hands, heads, and arms in a manner which is closely time-locked to the syntactic and semantic properties of what they are saying. Furthermore, McNeill notes, gesture and speech differ in their intrinsic representational characteristics and often serve different functions. For example, speech represents information in a highly structured, symbolic manner, whereas gestures represent information imagistically and holistically. Moreover, the primary function of speech is to describe things (objects, actions, events), whereas the function of gestures is to visually depict or highlight things.

[1] Because of the inherent verbal bias carried by the words "speaker" and "listener," the more apt words "communicator" and "addressee" will be used in the remainder of the paper.

Recent theoretical work by Clark (1996) lends support to McNeill's observations. Clark, using a semiotic typology proposed by the philosopher C. S. Peirce, proposes that language users have at their disposal several *methods of signaling—describing-as, demonstrating,* and *indicating.* These methods are all part of a single semiotic system. Language researchers have traditionally focused on the *describing-as* component of this system, which consists of abstract, conventionalized symbols emitted predominantly through the verbal channel. However, it must not be overlooked that communicators also demonstrate actions to addressees (e.g., "This is how you load the stapler," while physically demonstrating the action) and indicate objects in the speech environment using their eyes, hands, and/or demonstrative terms like "this" or "that" (as in, "That's the man I've been looking for," while pointing at the man across the room). These latter methods—demonstrating and indicating—involve nonverbal behaviors to a large degree. Utterances are typically a composite of these different semiotic types: "Demonstrating, indicating, and describing-as rarely occur in pure form. Just as most of Peirce's signs are 'mixed signs'—mixtures of icons, indices, and symbols—most signals are composite signals" (Clark, 1996, p. 161).

Thus, there are sound theoretical reasons to believe that nonverbal behaviors such as hand gestures play a significant role in communication. Indeed, recent studies have provided evidence that people are sensitive to information that is conveyed through gestures in multiple contexts (problem solving: Goldin-Meadow, Wein, & Chang, 1992; Kelly & Church, 1997, 1998; lexical discrimination: Thompson & Massaro, 1994; narrative processing: McNeill, Cassell, & McCullough, 1994). Given this, it seems surprising that psycholinguists have not also considered the role that gestures play in pragmatic processing. This lack of attention to nonverbal behaviors is particularly surprising in light of the classic observation in pragmatics that "speech underdetermines meaning." Traditionally, the move that theorists and researchers

have made in response to this observation has been to look outside the communicative act to determine meaning. However, valuable information about an utterance's meaning may be contained within the communicative act itself—in the nonverbal behaviors that frequently and naturally accompany speech.

The idea that a communicator's nonverbal behavior can, under certain circumstances, contribute to the meaning of an utterance is uncontroversial. Yet little is known about *how* these two sources of information are combined in comprehension and in memory. Do gesture and speech contribute to comprehension in an independent, additive manner, or do their meanings interact? Is gesture merely context for speech, or can speech also guide the interpretation of gesture? Do memory systems maintain separate traces for information conveyed verbally and nonverbally, or do people store both in an integrated trace? If people store gesture and speech as part of an integrated message, then information conveyed gesturally should have an impact on verbatim memory.

It is important at the outset to clarify the scope of the claims that we intend to make about gesture. Krauss and colleagues (Krauss, Morrel-Samuels, & Colasante, 1991) have noted that certain kinds of gestures are produced by communicators to facilitate language production (which we will refer to as "facilitative gestures"), rather than to communicate information to addressees. Thus, some forms of gestures may not be communicative, though this claim is in need of empirical support. For our purposes here, we focus on two kinds of gestural information which appear to be clearly intended as communicative, or "m-intended" (Grice, 1957). The first kind of gestures are manual pointing gestures, or deictic gestures, which serve to establish joint attention with an addressee. The second kind are iconic gestures, which imagistically depict objects, qualities, or activities. These two forms of gesture correspond to Clark's (1996) "indicating" and "demonstrating" methods of signaling, respectively. It is important to note, however, that they do not allow us to directly test claims about the impact of facilitative gestures on comprehension.

In the following four experiments, we look at the role that hand gestures play with speech in people's construal of pragmatic meaning during communication. Our first three experiments demonstrate that manual pointing gestures can contribute to the understanding of indirect requests, a pragmatically ambiguous speech act. In Experiment 1 we investigate the idea that pointing gestures contribute to the meaning of an indirect request. Experiment 2 provides an important control showing that only speech and gesture in combination yield the meanings of the indirect requests we used. Experiment 3 investigates how speech and gesture are combined. The final experiment, Experiment 4, extends the generality of our findings by examining other kinds of speech acts and other kinds of gestures.

## EXPERIMENT 1

This experiment examines the role of manual pointing gestures in understanding indirect requests. Indirect requests are a kind of pragmatically ambiguous utterance, because nothing in the "verbal" (i.e., spoken or written) portion of the message differentiates between a request and a declarative statement. For example, when someone says, "It's getting hot in here," an addressee must appeal to aspects of the communicative context in order to determine whether the speaker intended this as a request to do something, such as open a window, or only as a remark. Previous research suggests the importance of several factors in understanding indirect requests, specifically conversational implicature (Clark & Lucy, 1975), anticipating an addressee's "greatest potential obstacle" to compliance (Francik & Clark, 1985; Gibbs, 1986), conventionality (Gibbs, 1983), and the relative status of the speaker and addressee (Holtgraves, 1994).

However, the role of an important source of information about a communicator's intention—the communicator's gesture—has yet to be investigated. In line with Clark's observation that communicative acts often employ symbolic

and indexical methods of signaling, it is likely that when communicators make requests which involve an action on objects in the environment, they will occasionally indicate the relevance of these objects by pointing at them (e.g., looking at mother/child interactions: Shatz, 1978). To return to our example, when someone says, "It's getting hot in here," consider how you would understand this statement if the communicator was simultaneously pointing toward a closed window. The meaning is clear—the communicator is requesting that you open the window. But do people use this information or do they only pay attention to speech? We attempt to answer this question by examining people's understanding of indirect requests in the presence or absence of pointing. We predict that people will understand indirect requests more often in the presence of deictic gestures.

TABLE 1

*Scene*
Adam and Bill are returning home and meet in the street in front of their apartment.
Adam is on his bicycle, and Bill is walking.
*Dialogue*
Adam: Hey, did you get the burgers?
Bill: Oh no, I forgot!
Adam: Well, the guests are going to be here soon. You better go get the burgers.
*Target sentence*
*Bill: But the store is clear across town!*

| Experimental condition | Description |
|---|---|
| 1. Speech Only: | Bill makes normal eye contact with Adam and keeps his hands at his side. |
| 2. Speech + Gesture: | Bill points at Adam's bike. |

## *Methods*

*Participants.* Sixteen college undergraduates (8 males and 8 females) from the University of Chicago participated in the study for payment.

*Materials.* A videotape consisting of 12 scenarios was used as the experimental stimulus. All of the scenarios were composed of two professional actors acting out scripted interactions between two roommates, Adam and Bill. The scenarios were filmed in typical apartment settings (e.g., living room, kitchen, porch, front yard). Each scenario ended with a target sentence which could possibly be construed as an indirect request or a literal statement.

Each request encouraged action on a particular object in the environment, which we refer to as the target object. The materials are provided as Appendix A. An example of the verbal dialogue from one of the scenarios can be seen in Table 1, in which the character Bill is attempting to get Adam to lend him his bicycle.

As shown in Table 1, the target sentence could be delivered either with or without a point to the target object. Importantly, no mention of the target object was made in the speech. Thus, different information was conveyed in gesture and speech.

In the Speech Only condition, the actors made normal eye contact and kept their hands at their sides while the target sentence was delivered. In the Speech + Gesture condition, the communicator pointed at the target object while delivering the target sentence. The addressee in the latter case avoided highlighting the target object by first making eye contact with the communicator and then focusing on the communicator's extended finger.

While filming the vignettes, we attempted to make both conditions as similar as possible except for the gesture accompanying the target sentence. In the Speech + Gesture condition, we instructed the actors to introduce the behaviors while speaking and to attempt to perform them in the way that felt most "natural" to them. We ran a control study on 20 additional participants showing that the only reliable differences between the conditions were the gestures accompanying the indirect requests.

*Procedure.* We told participants that the study concerned how well people understand everyday social interactions. We informed them that they would be watching a sequence of videotaped scenarios about events in the everyday lives of two characters, Adam and Bill, and that they should pay close attention to the scenarios, because we would ask questions about the characters. We asked one question per scenario. The question asked participants to try to

predict how the person who had been addressed last in the interaction (just before the scenario ended) would react to what had been communicated to him. Recall that all the scenarios ended with indirect requests, so that responses to this question would likely reflect participants' understanding of indirect requests. We emphasized that there were no right or wrong answers; participants should simply write down their best guess of how the addressee would respond.

Each group watched one of four videotapes, which had the same sequence of vignettes while the order of conditions was varied. Each group of participants saw each vignette in just one of four conditions, only two of which are of interest here (Speech Only and Speech + Gesture). (We do not report the results from the other two conditions.)[2] Thus, each participant saw three scenarios in the Speech Only condition and three in the Speech + Gesture condition, though no participant saw the same vignette in more than one condition. Four stimulus tapes were created so that participants would view each scenario in only one condition. In all four versions, the order of the scenarios was always held constant, but the order in which participants received the experimental conditions was counterbalanced. The procedure lasted approximately 20 min.

*Coding.* After all of the data were collected, we coded each response to determine whether the respondent had interpreted the target sentence as an indirect request. We were primarily interested in whether participants understood the correct intention behind the indirect requests (though there were a number of possible types of responses, such as interpreting the utterances literally or attributing an incorrect intention to the request). Responses in which participants indicated that they understood the intention of the requests are referred to as Intended Action responses. Take the above "bike" vignette as an example. An Intended Action would be assigned if the participant, in response to the target sentence, said, "Adam will lend his bike to Bill." In this case, the participant correctly understands the specific intention behind the request by indicating the appropriate action in response to that request.

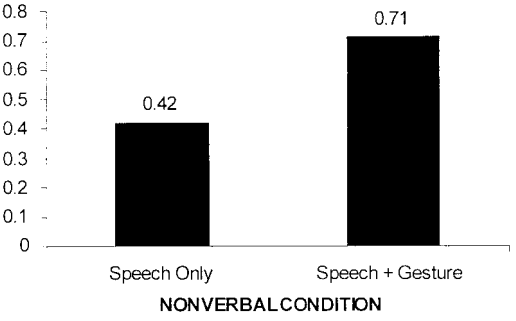


**FIG. 1.** Intended Action responses by nonverbal condition, Experiment 1.

*Results and Discussion*

Our prediction was that gestures accompanying an indirect request will increase the likelihood that participants will grasp the intended meaning of the request. To test this idea, we compared the percentage of times that subjects produced Intended Action responses in the Speech Only and the Speech + Gesture conditions. The results are summarized in Fig. 1. In the Speech Only condition, participants understood the intention of the request 42% of the time, and in the Speech + Gesture condition, participants understood the intention 71% of the time. A paired *t* test on the arcsine transformed values revealed a significant effect by subjects ($t_1(1,15) = 3.04$, $p < .05$) and by items ($t_2(1,11) = 2.89$, $p < .05$); see Fig. 1.

This result suggests that the deictic gestures used in this experiment make it easier to understand the specific intentions that underlie the

[2] We also manipulated the presence or absence of eye gaze toward the target object in our original experiment, which yielded four conditions: Speech Only, Speech + Gaze, Speech + Gesture, and Speech + Gaze + Gesture, but the resolution of the video medium made it difficult to determine the speaker's direction of gaze, so we have excluded the Speech + Gaze condition from our analysis. The analysis comparing Speech + Gesture and Speech + Gaze + Gesture revealed no significant differences, so we have opted to include only the Speech + Gaze + Gesture condition.

indirect requests. These data provide empirical support for Clark's semiotic typology, suggesting that speech and deictic gestures both play a role in the understanding of indirect requests. On the other hand, what we may have inadvertently done in our first experiment is to show that pointing gestures *by themselves* can constitute an indirect request. In other words, the results from this experiment are not sufficient to conclude that speech and gesture combine to create the indirect meaning. The following experiment introduces an important control by examining the individual contributions of speech and gesture to comprehension.

## EXPERIMENT 2

This experiment introduces a new condition to the Speech Only and Speech + Gesture conditions—a "Gesture Only" condition—which consists only of pointing gestures without speech. Using these three conditions, we attempt to control for the possibility that gestures by themselves do all the work in the Speech + Gesture condition of Experiment 1. The overall prediction is that people will better grasp the meaning of indirect requests when speech and pointing gestures are presented in combination versus when either is presented in isolation (i.e., Speech Only or Gesture Only).

### *Methods*

*Participants.* Eighteen University of Chicago undergraduates (9 males and 9 females) were paid to participate in the study.

*Materials.* Our materials were identical to those of Experiment 1: we used the same videotapes with the same scenarios. The Gesture Only condition was created by muting the audio playback of the Speech + Gesture condition during the target utterance. Table 2 shows an example and conditions.

Notice that in this example, there were three forms of the target sentence. In the Speech Only condition**,** the actor delivered the target sentence, making normal eye contact and keeping his hands at his side. In the Speech + Gesture condition, the actor delivered the target sentence while pointing at the open screen door. These first two conditions are identical to the two conditions used in Experiment 1. Finally, in the Gesture Only condition, the actor's speech was muted while he pointed and delivered the target utterance. In all other respects, the scenario was identical to the Speech + Gesture condition. In all, there were four instances of each of the three conditions, yielding a total of 12 vignettes. Other than the different types of conditions, the design of the tape was identical to Experiment 1.

TABLE 2

*Scene*
Bill is sunning himself on the porch. After Bill swats at a couple of flies, Adam (who is inside) opens the screen door (which he does not close) and enters the porch area.
*Dialogue*
Adam: I found that book I was looking for. (Pause) Man, it's hot out here.
*Target sentence*
*Bill: Yeah, and the flies are out.*

| Experimental condition | Description |
|---|---|
| 1. Speech Only: | Bill makes normal eye contact and keeps his hands at his side. |
| 2. Speech + Gesture: | Bill points at the open screen door. |
| 3. Gesture Only: | The audio portion of the target sentence is muted, and Bill points at the open screen door. |

*Apparatus.* A computer-controlled Sony Hi-8 VCR (Model EVO-8650) was used to play the videotape. During the playback of stimulus items, a PC computer continuously read the time code from the videotape. Upon reaching the time code corresponding to the beginning of the critical utterance, it sent a command to the VCR which muted the audio playback for the gesture only conditions.

*Procedure.* The instructions given to participants were the same as those in the first experiment. Once again, the participants' task was to write down how they thought that the addressee (the person spoken to last) would respond. The

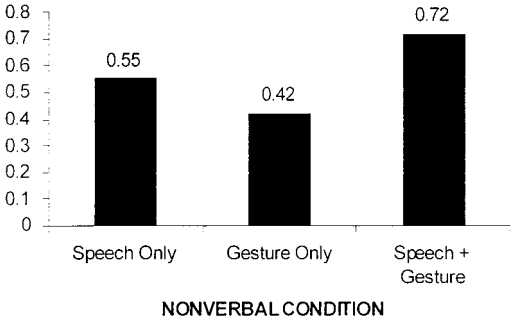


**FIG. 2.** Intended Action responses by nonverbal condition, Experiment 2.

procedure lasted approximately 20 min. The coding was identical to that of Experiment 1.

*Results and Discussion*

Our main prediction for this experiment was that participants would produce Intended Action responses more often in the Speech + Gesture condition than in either the Speech Only or Gesture Only condition. The results supported this prediction. In the Speech + Gesture condition, participants produced Intended Action responses 72% of the time, whereas in the Speech Only condition 55% of the time, and in Gesture Only 42% of the time. We submitted the arcsine transformed values of the data to two separate one-way repeated measures ANOVAs, which revealed a significant effect both by subjects ($F_1(2,51) = 7.18$, $p < 0.001$ Greenhouse-Geiser $\epsilon$) and by items ($F_2(2,33) = 7.34$, $p < 0.05$ Greenhouse-Geiser $\epsilon$). Additionally, planned comparisons revealed reliable differences between Speech Only and the Speech + Gesture conditions ($t(1,17) = 2.11$, $p < .05$), as well as between the Gesture Only and the Speech + Gesture conditions ($t(1,17) = 3.82$, $p < .005$). However, there was no difference between the Speech Only and the Gesture Only conditions ($t(1,17) = .15$, ns). The results are summarized in Fig. 2.

If the gesture by itself was driving the effect in the Speech + Gesture condition, then we would expect no significant differences between the Speech + Gesture and the Gesture Only conditions. However, participants in the Speech + Gesture condition were far more likely to interpret the action as an indirect request than those in the Gesture Only condition. Thus, neither speech nor gesture alone reveals the meaning of indirect requests as well as the combination of the two pieces together.

The results from the first two experiments convincingly demonstrate that speech and gesture combine to determine meaning of indirect requests. But the question of *how* they combine remains unresolved. The view implicit in many pragmatic theories is that gesture and speech are combined additively in the conveyed meaning of an utterance. In most theories of pragmatics, the intended meaning of an utterance is derived by assessing the literal meaning of speech in the pragmatic context in which it is uttered (Grice, 1975; Searle, 1975). The communicator's nonverbal behavior is considered to be part of this context. Although both sources of information would ultimately figure in the conveyed meaning, these theories suggest that the meaning of speech and the meaning of gesture are computed independently of one another. To return to our example, when the communicator says, "The flies are out," and points to the open screen door, the addressee interprets the statement about the flies and independently notes that the speaker is pointing at the door. The addressee would then combine these two meanings to grasp the communicator's request to close the door. We call this hypothesis the *additive contribution hypothesis.*

On the other hand, it is possible that verbal and nonverbal information form an integrated message for the addressee. In contrast to the additive contribution hypothesis, this *interactive contribution hypothesis* states that the meanings of speech and gesture are computed interactively: that speech is context for gesture just as gesture is context for speech.[3] Thus, these two hypotheses make different predictions about whether speech will influence the interpretation of gesture. In our example, the additive contribution hypothesis

[3] We thank Sotaro Kita for first suggesting this idea to us.

predicts that understanding one piece of information (e.g., the speech) should be independent of understanding the second piece of information (e.g., gesture). In contrast, the interactive contribution hypothesis predicts that understanding one piece of information will be affected by understanding the other. For example, people should understand the meaning of "The flies are out" differently with and without an accompanying gesture to the screen door, and people should understand the referent of the gesture differently with and without the accompanying speech.

A posthoc analysis of the results in Experiment 2 seems to support the additive contribution hypothesis. Consider the percentage of times that people failed to understand the meaning of indirect requests in the Speech Only and Gesture Only conditions. People did not understand the meaning 45% of the time in the Speech Only condition and 58% of the time in the Gesture Only condition. When these two probabilities are combined in an independent fashion (the additive model), the outcome is a 26% chance of failing to understand the indirect requests. This 26% is practically identical to the percentage of times that people did not understand the meaning when the two pieces of information were both present in the Speech + Gesture condition (29% of the time). Thus, this superficial analysis suggests that speech and gesture contribute in an additive fashion to the meaning of indirect requests.

In Experiment 3, we attempt to more directly test between the two hypotheses. If speech and gesture are related in an additive fashion, then understanding the meaning of one piece of information should not be influenced by understanding the other piece of information. For example, when one of our actors points to an object, participants should be able to identify the referent of that gesture equally well in the presence or the absence of accompanying speech. On the other hand, if speech and gesture are related in an interactive fashion, participants should have a different understanding of a gesture when no speech accompanies it.

## EXPERIMENT 3

Experiment 3 asks how spoken information influences the interpretation of pointing gestures. Specifically, we compare people's ability to identify referents of manual pointing gestures when accompanied by speech to when the speech is muted. In this experiment, as in the previous two, the speech did not mention the referent of the pointing gesture, but contained distinct information. Thus, if participants are better at identifying the referents of pointing gestures when accompanied by speech than when presented without, this supports the interactive contribution hypothesis.

### *Methods*

*Participants.* Fifteen University of Chicago students (7 males and 8 females) participated in the experiment for payment.

*Materials.* We used the same videotapes that were used in Experiments 1 and 2. We were interested only in two conditions: Speech + Gesture and Gesture Only. The Gesture Only condition was exactly identical to the Speech + Gesture condition, except that during the critical utterance, the audio was muted.

*Apparatus.* A computer-controlled Sony Hi-8 VCR (Model EVO-8650) was used to play the videotape. A PC computer, configured to control the Hi-8, muted the audio playback during the critical utterance of the gesture only condition.

*Procedure.* We told participants that the experiment concerned how well people could understand nonverbal pointing behaviors. Specifically, we told them that they would be watching 12 video segments in which they would be seeing someone pointing at objects. After participants viewed each segment, they were asked to identify the referred-to object. This identification was made in two ways. First, participants were asked the open-ended question, "What object did the person point at?" Then, participants were given a forced-choice task requiring them to choose from five potentially referred-to items.

Participants watched one of two videotapes,

each videotape having the same sequence of vignettes in the same order, but in different experimental conditions. The two conditions alternated within each videotape. Furthermore, each scenario appeared only once on any given videotape, either in the Speech + Gesture or in the Gesture Only condition. Participants saw 6 segments in each condition, for a total of 12. The procedure lasted approximately 15 min.

*Coding.* The coding for the open-ended questions was straightforward. If participants wrote down the intended object of the indirect request, they were assigned an Intended Object code. For example, in the "flies" scenario, the intended object was the open screen door. The forced-choice task worked the same way; if participants circled the appropriate choice, they were assigned an Intended Object code for that vignette.

*Results and Discussion*

The general prediction for both measures was that participants would select the intended object more often in the Speech + Gesture condition than in the Gesture Only condition. For the open-ended question, participants produced Intended Object responses, on average, 67% of the time for the Gesture Only condition and 91% of the time for the Speech + Gesture condition. A paired $t$ test on the arcsine transformed data revealed a significant effect both by subjects ($t(1,14) = 6.99$, $p < .001$) and by items ($t(1,11) = 3.74$, $p < .001$). For the forced-choice question, participants produced Intended Object responses, on average, 63% of the time for the Gesture Only condition and 89% of the time for the Speech + Gesture condition. A paired $t$-test on the arcsine transformed data revealed a significant effect for both subjects ($t(1,14) = 6.78$, $p < .001$) and for items ($t(1,11) = 3.71$, $p < .001$); see Fig. 3.

This experiment demonstrates that the referent of the pointing gestures in our scenarios is determined, in part, by the speech that accompanies them. This finding allows us to reject the idea that speech and gesture contribute to meaning in a strictly additive fashion. This makes sense when one considers that pointing is a communicative act which is itself potentially ambiguous. When Bill points at the bike, for example, it may be quite unclear to an addressee what he is pointing at—he could be pointing at the sidewalk, the ground, the bike, or anything else that happened to intersect the axis extending from the end of his finger. Thus, whatever extra information we have about Bill's intention could be used to constrain the set of possible referents. The results from Experiment 3 suggest that the *speech itself* provides this important constraint.

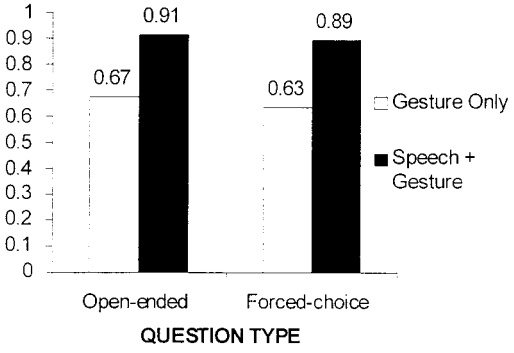


**FIG. 3.** Open-ended and forced-choice responses by nonverbal condition, Experiment 3.

To summarize briefly, in accordance with Clark's semiotic typology, Experiments 1 and 2 demonstrate that pointing gestures can contribute to the meaning of indirect requests. Experiment 3 provides evidence against the additive contribution model and in support of the interactive contribution model. In the final experiment, Experiment 4, we attempt to generalize our findings in several ways. In Experiments 1–3, we tested only two components of Clark's semiotic typology: *describing-as* and *indicating.* In Experiment 4, we extend our testing of Clark's typology by investigating the impact that *demonstrative* methods of signaling have on people's understanding of speech. To do this, we introduce a new type of gesture: iconic gestures. Moreover, we look at the role that these gestures play in different types of communicative acts: descriptions of activities and events. Finally, we wanted to test construal of gesture in a different way. Using a memory paradigm,

we investigated whether people had difficulty monitoring the source of information conveyed through spoken and gestured channels.

## EXPERIMENT 4

Iconic, or representational, gestures (McNeill, 1992) are hand gestures which represent information imagistically—depicting such things as object attributes, actions, and spatial relationships. It is well documented that iconic gestures are abundant in a wide range of communicative settings (Church et al., 1995; Clark, 1996; Goldin-Meadow, Alibali, & Church, 1993; McNeill, 1992). Moreover, these gestures often contribute to an utterance's intended meaning, making them good candidates for investigating the role that gesture plays in pragmatic comprehension. For example, suppose a distraught motorist were to explain how his car had been hit by another car: in speech, the person could say, "I didn't see it coming," while gesturing the image of another car blind-siding his car from the side. In this way, gesture may serve the pragmatic function of revealing the attributes of the cars, the direction of movement of the cars, and the spatial relationship of the cars. Thus, the combination of speech and gesture may send a more clear and thorough picture of what the communicators intend to communicate.

In the following experiment, we use a memory paradigm to test the idea that information conveyed through iconic gestures is incorporated into the intended meaning of a message. If people cannot help but include gestural information in their recall for speech, this would be solid evidence that iconic gestures play an significant role in determining the intended meaning of an utterance.

### *Methods*

*Participants.* Fifteen Northeastern Illinois University college undergraduates (8 males and 7 females) participated in the study for course credit.

*Materials.* The experimental video stimulus consisted of a woman (a professional actress) making 10 isolated statements about everyday situations. Each statement was one sentence long and referred to some sort of ordinary activity. The materials are provided as Appendix B. Half of the statements were made with speech alone, and the other half were accompanied by gesture. For an example, refer to Table 3.

TABLE 3

*Target sentence*
*My brother went to the gym.*

| Experimental condition | Description |
|---|---|
| 1. Speech Only: | The woman makes no accompanying gestures. |
| 2. Speech + Gesture: | The woman makes a gesture depicting the shooting of a basketball. |

As shown in Table 3, there were two experimental conditions. In the Speech Only condition, the woman made statements with no accompanying gesture (i.e., her hands were at her side). In the Speech + Gesture condition, the woman made the statement but also produced iconic gestures along with the speech. All of the gestures provided information that was not expressed in the speech.

While filming the vignettes, we attempted to make the two conditions as similar as possible except for the gesture accompanying the statement. In the Speech + Gesture condition, we instructed the actress to perform the gestures in a way that felt most "natural" to her.

*Procedure.* Participants were told that they were going to watch videotaped segments of a person recounting everyday events. They were informed that they should pay close attention because after they viewed all of the segments, they would be asked to recall what the woman had said in the segments. Participants then watched all 10 of the segments.

Immediately following this task, we collected data using a cued recall procedure. Participants were given written prompts to probe their memory for what was said. For example, the prompt for the above statement was, "The woman talked about her brother; what did she say?" Participants were urged to try to write down the

*exact* words that the woman had said. The entire procedure—stimulus presentation and recall—lasted approximately 15 min.

*Coding.* We were primarily interested in two types of response from the participants. First, we looked at how well participants remembered the spoken portions of the messages. Specifically, we looked at the extent to which they recalled the speech verbatim (e.g., after seeing the "basketball" vignette, if a person recalled, "My brother went to the gym") or the extent to which they recalled the gist of the speech (e.g., "My brother left for the gymnasium"). Second, we looked at whether participants misremembered the spoken portion of the indirect request and instead, "remembered" the intended meaning of the request. Because these recollections could be traced back to the gesture, they were coded as "Traceable Additions." So in the above "basketball" example, a participant would be assigned an Traceable Addition code if she misremembered the woman as having said, "My brother went to play basketball."

### *Results and Discussion*

Our prediction was that people would incorporate gestural information into their memories for speech in the Speech + Gesture condition. To determine this, we compared the percentage of Traceable Additions produced in the Speech + Gesture condition to the percentage of times additions were produced in the Speech Only condition. A paired *t* test analysis on the arcsine transformed data revealed that participants produced significantly more Traceable Additions in the Speech + Gesture condition (23% of the time) than in the Speech Only condition (0% of the time) both by subjects ($t_1(1,14) = 3.48$, $p < .001$) and by items ($t_2(1,9) = 4.1$, $p < .001$); see Fig. 4. (Refer to Appendix C for all of the Traceable Additions produced in response to Speech + Gesture stimuli.)

The second analysis compared how well people remembered just the spoken portion of the message in both conditions, regardless of whether gesture intruded. We found that the quality of the memory for speech was influenced by the presence or the absence of gesture. Interestingly, participants produced a combined percentage of verbatim and gist repetitions 58% of the time in the Speech Only condition and 77% of the time in the Speech + Gesture condition, which was significant by subjects ($t(1,14) = 1.99$, $p < .05$) but not by items ($t(1,9) = .97$, $p = .15$, ns). This trend (though nonsignificant by items) suggests a possible effect of gesture on memory for speech, which would appear to be in line with the interactive contribution hypothesis.

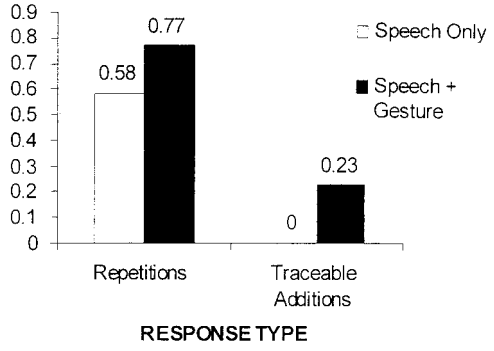


**FIG. 4.** Repetitions and Traceable Additions by nonverbal condition, Experiment 4.

The above findings provide convincing evidence that information conveyed through iconic gesture is incorporated in what participants consider an utterance's intended meaning. These results are particularly striking in light of the strict instructions to recall just the spoken information—information that, in principle, could be easily understood independently of the accompanying gesture. Moreover, when probed in an exit interview, participants rarely remembered having received the information through gesture. Though more research is needed, the fact that participants were not good at monitoring the source of information in the videotapes suggests that gesture and speech may be tightly linked in comprehension (for a review of the source monitoring literature, see Johnson, Hashtroudi, & Lindsay, 1993).

## GENERAL DISCUSSION

To summarize, our experiments demonstrate that certain nonverbal behaviors, such as deictic

and iconic gestures, can have a powerful impact on how people comprehend and remember pragmatic communication. Experiment 1 showed that the presence of pointing gestures made respondents more likely to interpret utterances as indirect requests than when they only heard speech. Experiment 2 replicated Experiment 1 and provided an important control for the possibility that the difference in Experiment 1 was attributable to gesture alone. Experiment 3 rejected the additive contribution hypothesis of speech and gesture processing in favor of the interactive contribution hypothesis, by showing that speech often constrains the meaning of gesture. Finally, Experiment 4 extended the findings of Experiments 1–3 to include different types of gestures and different kinds of speech acts. Taken together, these experiments suggest that speech and gesture may interact to codetermine meaning in communication.

Even though the results from Experiment 3 reject the strictly additive mdel of interaction, we cannot conclusively determine just *how* speech and gesture interact as comprehension unfolds. However, it does seem that in oncine comprehension, understanding gesture may be affected by understanding speech and vice versa. In other words, the experiment suggests that speech and gesture may (at least some of the time) interactively contribute to the meaning of a communicative act. This claim receives support from qualitative analyses of our items. For example, in the "flies" scenario, a point to an open screen door without the corresponding speech, "The flies are out," elicited responses such as "Do you want to go inside?" And when speech was presented without gesture, participants also had difficulty interpreting the meaning, indicating that the addressee would respond by saying such things as, "Yeah, the flies are bad this year," or "Did you get bitten?" However, with information from both modalities instead of just one, people were much more likely to correctly understand the intended meaning. Examples such as these suggest that not only does gesture disambiguate the meaning of speech, but speech disambiguates the meaning of gesture.

One way to definitively test the idea that speech and gesture have an interactive relationship is to employ an online methodology. Previous researchers have used online methodologies (e.g., sentence verification tasks) to show that people initially integrate contextual information into their understanding of sarcasm (Gibbs, 1986) and indirect requests (Holtgraves, 1994). These researchers have shown that people do not first encode the literal meaning of an utterance and then subsequently use context to understand the intended meaning. Rather, they have shown that people can initially bypass the literal meaning of utterances and immediately grasp the intended meaning (Gibbs, 1979, 1983). We are currently designing online studies to test the idea that gestural information is initially integrated into the meaning of speech.

As a research program, the study of the pragmatic function of nonverbal information could have potentially broad theoretical and methodological implications. From its foundations in the philosophical literature to present-day psycholinguistic research, the field of pragmatics has taken as its point of departure the spoken word (or perhaps more correctly, the written word). This rarefied view of communication leads us to look for pragmatic meaning outside of the face-to-face conditions which constitute the primary arena of human communication. As a consequence, the traditional pragmatic problem that "speech underdetermines meaning" might be overstated, simply because an important source of pragmatic information—nonverbal behavior—has not been considered. By expanding the linguistic unit of analysis to include information conveyed through a communicator's eyes, hands, or tone of voice—which, along with speech, seem to actively codetermine the meaning of an utterance—psycholinguists may begin to approach the study of language from a perspective that is more in line with what happens in everyday, face-to-face communication.

In addition to these theoretical implications, we believe that the present study has implications for the methods with which psycholinguists approach the study of pragmatics. The

overwhelming majority of research on how people comprehend language employs text-based methodologies to answer various questions of interest. For example, participants are given brief text passages which describe situations and are asked to indicate how characters in the narratives will understand utterances produced in those situations. Indeed, studies employing text-based methodologies have generated many valuable insights into how people process and understand pragmatic information when reading texts. It is not clear, however, just how well the results from these studies generalize to face-to-face interactions between people in everyday life. There are good reasons to believe, at least for some communicative phenomena, that text-based methodologies may not be the best way to study what happens in face-to-face interactions. The video methodology which we used in these experiments allowed participants to both hear *and see* people interacting with each other. Because of the more realistic setting, the data from the video methodology may be more representative of what happens in real face-to-face interactions.

However, a caveat is in order. Even though the use of the video methodology allows for a more realistic investigation of what happens in face-to-face interactions, it is not, of course, an investigation of what does happen in real life, face-to-face interactions. Just like text-based studies, it is an approximation—albeit a closer approximation, we argue—of what happens in real life. As noted in the introduction, because we used nonspontaneous gestures that are clearly communicative or "m-intended," these experiments do not allow us to directly test claims about the impact of spontaneously produced, facilitative gestures on pragmatic comprehension and memory. Additionally, one might question the applicability of our findings by arguing that the gestures used in our stimuli were artificial and contrived. Yet studies looking at naturalistic detection of unplanned, spontaneously produced gestures in other contexts (Kelly & Church, 1998; Thompson & Massaro, 1994) show that people are quite sensitive to the information conveyed through deictic and iconic gestures. It would be surprising if this natural sensitivity did not also hold for pragmatic communication in everyday situations.

To return to our original problem: How do we understand what someone means by what they say? Most of the research on pragmatics has assumed that what is "said" is nothing more than the words that are spoken or written. However, we agree with Clark (1996) that researchers need to take a broader view of communication and include, for example, information from other modalities such as hand gestures. As we have shown, these behaviors make a substantial contribution to an utterance's meaning. Given this, we hope that future research on pragmatics will begin to take hand gestures seriously and strongly consider the role that these pervasive and important behaviors play along with speech in communication.

## APPENDIX A: The 12 Indirect Request Scenarios, Experiments 1–3

*Scene 1:* Adam and Bill in front of their apartment. Adam is on his bicycle.

Adam: Hey, did you get the burgers?

Bill: Oh no, I forgot!

Adam: Well, the guests are going to be here soon. You better go get the burgers.

*Bill's indirect request*

Speech Only: But the store is clear across town.

Speech + Gesture: But the store is clear across town (points at bicycle).

*Scene 2:* Adam and Bill are in the sun room. Bill is working on a laptop.

Bill: So we got the introduction done. Do you want to take a break?

Adam: No. Let's work another hour (walks toward the window and opens it).

*Bill's indirect request*

Speech Only: But I'm getting cold.

Speech + Gesture: But I'm getting cold (points at open window).

*Scene 3:* Bill is eating a sandwich in the living room. Adam has finished his.

Adam: That was a great sandwich.

Bill: You're done already. I've never seen anyone eat as much as you!

*Adam's indirect request*

Speech Only: Actually, I'm still pretty hungry.

Speech + Gesture: Actually, I'm still pretty hungry (points at Bill's sandwich).

*Scene 4:* Bill is drinking a beer in front of the TV. Adam enters with his own beer.

Adam: Did I miss anything? (attempting to twist off beer cap).

Bill: No. It's still on pause.*Adam's indirect request*

Speech Only: This isn't a twist-off.

Speech + Gesture: This isn't a twist-off (points at bottle opener on table).

*Scene 5:* Bill is lying on his bed in his room. Adam enters.

Adam: Hey, did I wake you?

Bill: No, I was just resting. You got some mail (hands Adam a postcard).

*Adam's indirect request*

Speech Only: I can't read this.

Speech + Gesture: I can't read this (points at lamp next to bed).

*Scene 6:* Adam and Bill just finished watching a movie in the living room.

Bill: I really picked a winner this time (removes the movie from the VCR).

Adam: Not only was it bad, it was so long (looks at watch). Well, I've got to head out.

*Bill's indirect request*

Speech Only: Are you in a hurry?

Speech + Gesture: Are you in a hurry (points at video cassette)?

*Scene 7:* Adam is taking down a picture from the wall. Bill is hammering off camera.

Adam: We've got to move this one again.

Bill: (Coming into view holding hammer) I thought we said that it looked alright there.

Adam: (trying to pull out nail) It's too high.

*Adam's indirect request*

Speech Only: I can't get this nail out.

Speech + Gesture: I can't get this nail out (points at hammer).

*Scene 8:* Adam is watching TV. There is a big mess on the table.

Bill: So I see that you had friends over last night (looks around room).

Adam: Yeah. It was a good time.

Bill: I had an alright time with my parents last night.

Adam: Oh. Are they still in town?

*Bill's indirect request*

Speech Only: Actually, they're going to be here any minute.

Speech + Gesture: Actually, they're going to be here any minute (points at mess).

*Scene 9:* Adam and Bill are eating dinner at the table.

Adam: Be careful of the chili peppers. They're really hot.

Bill: Alright (pauses and takes a bite). What should we do later?

Adam: I was thinking we should go to that party (pours water from a pitcher to his glass).

*Bill's indirect request*

Speech Only: You're right—those peppers are hot.

Speech + Gesture: You're right—those peppers are hot (points at pitcher of water).

*Scene 10:* Adam is in the kitchen washing dishes. Bill enters with a backpack on.

Bill: Wow, we really made a mess last night, didn't we?

Adam: Yeah (notices backpack). Where are you off to?

Bill: I have to go run some errands.

Adam: So are you going to have time to help clean up?

Bill: Yeah, but I really have to do these things now. I'll be back in a few hours.

*Adam's indirect request*

Speech Only: Are you going out the back door?

Speech + Gesture: Are you going out the back door (points at overflowing garbage)?

*Scene 11:* Bill is preparing dinner. Adam enters.

Adam: Smells good in here.

Bill: Thanks. How's the game going?

Adam: The Bulls are rocking!

Bill: Are people getting hungry?

Adam: Yeah, I think so.

*Bill's indirect request*

Speech Only: Actually, it's almost ready.

Speech + Gesture: Actually, it's almost ready (points at placemats and a stack of plates).

*Scene 12:* Bill is sunning himself on the porch. Adam enters the porch area.

Adam: I found that book I was looking for. (Pause) Man, it's hot out here.

*Bill's indirect request*

Speech Only: Yeah, and the flies are out.

Speech + Gesture: Yeah, and the flies are out (points at open screen door).

## APPENDIX B: The 10 Iconic Gesture Segments, Experiment 4

*Segment 1*

Speech: It was bad in the room.

Gesture: Waves hand in from of nose to indicate bad smell.

*Segment 2*

Speech: I told my friend about the party.

Gesture: Places hand next to ear to indicate talking on the phone.

*Segment 3*

Speech: The weight lifter was out of shape.

Gesture: Extends both hands from stomach to indicate fatness.

*Segment 4*

Speech: My brother went to the gym.

Gesture: Makes shooting gesture to indicate playing basketball.

*Segment 5*

Speech: The church is around the corner.

Gesture: Makes turning gesture to indicate a right hand turn.

*Segment 6*

Speech: The stockbroker was up late last night at the restaurant.

Gesture: Makes gesture near mouth to indicate drinking.

*Segment 7*

Speech: The carpenter was working in the garage.

Gesture: Moves hand up and down to indicate hammering.

*Segment 8*

Speech: The lawyer got ready for work.

Gesture: Moves hand in front of mouth to indicate brushing teeth.

*Segment 9*

Speech: The camper caught a fish.

Gesture: Holds hands far apart to indicate a very large fish.

*Segment 10*

Speech: The cook stepped outside for a minute.

Gesture: Makes a gesture near her mouth to indicate smoking a cigarette.

## APPENDIX C: The 17 Traceable Additions, Experiment 4

1. She talked about a room, what did she say?
   Subject 2: "Bad smell."
   Subject 4: "It smelled."
2. She talked about her best friend, what did she say?
   No data.
3. She talked about a weight lifter, what did she say?
   Subject 4: "He had a gut."
4. She talked about her brother, what did she say?
   Subject 13: "Went to play ball."
5. She talked about a church, what did she say?
   Subject 7: "The church was that way to the right."
   Subject 9: "It's this way to the right."
   Subject 10: "The church is over there to her right."
6. She talked about a stockbroker, what did she say?
   Subject 9: "He was at the restaurant drinking."
   Subject 13: "Up late drinking."
   Subject 14: "She said the stockbroker was up late last night drinking."
7. She talked about a carpenter, what did she say?
   Subject 1: "Hammering."
   Subject 5: "He hammered something."
8. She talked about a lawyer, what did she say?
   Subject 1: "Brushing his teeth."
9. She talked about a camper, what did she say?
   Subject 13: "He caught a tiny fish."
   Subject 14: "He caught a fish about 16 inches big."
10. She talked about a cook, what did she say?
    Subject 13: "Went out for a smoke."
    Subject 14: "She said that he went outside for a smoke."

## REFERENCES


Alibali, M., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology,* **25,** 468–523.

Argyle, M. (1973). The syntaxes of bodily communication. *International Journal of Psycholinguistics,* **2,** 71–91.

Austin, J. L. (1962). *How to do things with words.* Cambridge, MA: Harvard Univ. Press.

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology,* **29(5),** 832–843.

Bruner, J. (1984). Interaction, communication, and self. *Journal of the American Academy of Child Psychiatry,* **23(1),** 1–7.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition,* **23,** 43–71.

Church, R. B., Schonert-Reichl, K., Goodman, N., Kelly, S. D., & Ayman-Nolley, S. (1995). The role of gesture and speech communication as reflection of cognitive understanding. *Journal of Contemporary Legal Issues,* **6,** 123–154.

Clark, H. H. (1996). *Using language.* Cambridge, GB: Cambridge Univ. Press.

Clark, H. H., & Lucy, P. (1975). Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior,* **14,** 56–72.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, I. Sag, & B. Webber (Eds.), *Linguistic structure and discourse setting.* Cambridge: Cambridge Univ. Press.

Francik, E. P., & Clark, H. H. (1985). How to make requests that overcome obstacles to compliance. *Journal of Memory and Language,* **24,** 560–568.

Gibbs, R. W. (1979). Contextual effects in understanding indirect requests. *Discourse Processes,* **2,** 1–10.

Gibbs, R. W. (1983). Do people always process the literal meanings of indirect requests? *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **9,** 524–533.

Gibbs, R. W. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General,* **115,** 3–15.

Goldin-Meadow, S., Alibali, M., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review,* **100,** 279–297.

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction,* **9,** 201–219.

Grice, H. P. (1957). Meaning. *Philosophical Review,* **66,** 377–388.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and semantics,* (Vol. 3). New York: Academic Press.

Holtgraves, T. (1994). Communication in context: Effects of speaker status on the comprehension of indirect requests. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **20,** 1205–1218.

Johnson, M. K., Hashtroudi, S., & Lindsay, S. D. (1993). Source monitoring. *Psychological Bulletin,* **114,** 3–28.

Kelly, S. D., & Church, R. B. (1997). Can children detect conceptual information conveyed through other chil-

dren's nonverbal behaviors? *Cognition and Instruction,* **15(1),** 107–134.

Kelly, S. D., & Church, R. B. (1998). A comparison between children's and adults' ability to detect conceptual information conveyed through representational gestures. *Child Development,* **69,** 85–93.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology,* **61,** 743–754.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review,* **92,** 350–371.

McNeill, D. (1987). *Psycholinguistics: A new approach.* New York: Harper & Row.

McNeill, D. (1992). *Hand and mind: What gesture reveals about thought.* Chicago, IL: Univ. of Chicago Press.

McNeill, D., Cassell, J., & McCullough, K. E. (1994). Communicative effects of speech-mismatched gestures. *Research on Language and Social Interaction,* **27**(3), 223–237.

Searle, J. R. (1975). What is a speech act? In M. Black (Ed.), *Philosophy in America.* Ithaca: Cornell Univ. Press.

Shatz, M. (1978). On the development of communicative understandings: An early strategy for interpreting and responding to messages. *Cognitive Psychology,* **10,** 271–301.

Thompson, L. A., & Massaro, D. W. (1994). Children's integration of speech and pointing gestures in comprehension. *Journal of Experimental Child Psychology,* **57,** 327–354.

Tomasello, M. (1992). The social bases of language acquisition, *Social Development,* **1,** 67–87.

(Received August 13, 1998)
(Revision received January 4, 1999)